

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>7</sup> :

G10L 11/04

A1

(11) International Publication Number:

WO 00/25298

(43) International Publication Date:

4 May 2000 (04.05.00)

(21) International Application Number: PCT/CA99/01008

(22) International Filing Date: 27 October 1999 (27.10.99)

(30) Priority Data:

2,252,170

27 October 1998 (27.10.98)

CA

(71) Applicant (for all designated States except US): VOICEAGE CORPORATION [CA/CA]; Suite 200, 750, Chemin Lucerne, Ville Mont-Royal, Quebec H3R 2H6 (CA).

(72) Inventors; and

(75) Inventors/Applicants (for US only): BESSETTE, Bruno [CA/CA]; 1546 Pérodeau, Rock Forest, Quebec J1N 1L2 (CA). SALAMI, Redwan [CA/CA]; 963, Léo Laliberté, Sherbrooke, Quebec J1J 4L3 (CA). LEFEBVRE, Roch [CA/CA]; 259, avenue de la Bourgade, Canton de Magog, Quebec J1K 5R9 (CA).

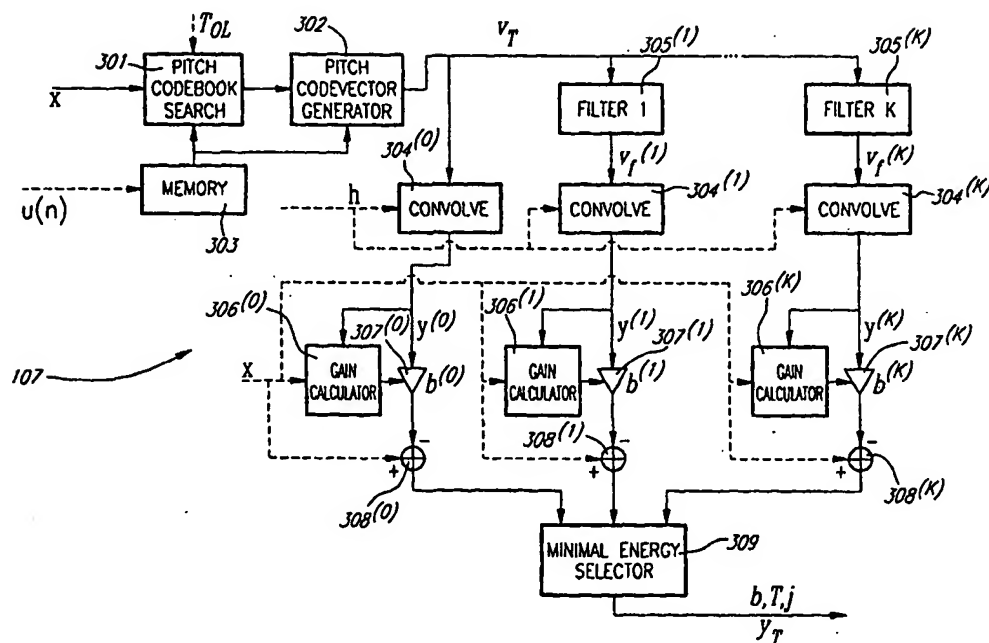
(74) Agents: DUBUC, Jean, H. et al.; Goudreau Gage Dubuc &amp; Martineau Walker, The Stock Exchange Tower, Suite 3400, 800 Place Victoria, Montreal, Quebec H4Z 1E9 (CA).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

(54) Title: A METHOD AND DEVICE FOR ADAPTIVE BANDWIDTH PITCH SEARCH IN CODING WIDEBAND SIGNALS



## (57) Abstract

An improved pitch search method and device for digitally encoding a wideband signal, in particular but not exclusively a speech signal, in view of transmitting, or storing, and synthesizing this wideband sound signal. The new method and device which achieve efficient modeling of the harmonic structure of the speech spectrum uses several forms of low pass filters applied to a pitch codevector, the one yielding higher prediction gain (i.e. the lowest pitch prediction error) is selected and the associated pitch codebook parameters are forwarded.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## A METHOD AND DEVICE FOR ADAPTIVE BANDWIDTH

### PITCH SEARCH IN CODING WIDEBAND SIGNALS

5

#### BACKGROUND OF THE INVENTION

##### 1. Field of the invention:

10

The present invention relates to an efficient technique for digitally encoding a wideband signal, in particular but not exclusively a speech signal, in view of transmitting, or storing, and synthesizing this wideband sound signal. More specifically, this invention deals with an improved pitch search device and method.

15

##### 2. Brief description of the prior art:

20

The demand for efficient digital wideband speech/audio encoding techniques with a good subjective quality/bit rate trade-off is increasing for numerous applications such as audio/video teleconferencing, multimedia, and wireless applications, as well as Internet and packet network applications. Until recently, telephone bandwidths filtered in the range 200-3400 Hz were mainly used in speech coding applications. However, there is an increasing demand for wideband speech applications in order to increase the intelligibility and

25

naturalness of the speech signals. A bandwidth in the range 50-7000 Hz was found sufficient for delivering a face-to-face speech quality. For audio signals, this range gives an acceptable audio quality, but still lower than the CD quality which operates on the range 20-20000 Hz.

5                   A speech encoder converts a speech signal into a digital bitstream which is transmitted over a communication channel (or stored in a storage medium). The speech signal is digitized (sampled and quantized with usually 16-bits per sample) and the speech encoder has the role of representing these digital samples with a smaller number of  
10                   bits while maintaining a good subjective speech quality. The speech decoder or synthesizer operates on the transmitted or stored bit stream and converts it back to a sound signal.

                  One of the best prior art techniques capable of achieving a  
15                   good quality/bit rate trade-off is the so-called Code Excited Linear Prediction (CELP) technique. According to this technique, the sampled speech signal is processed in successive blocks of  $L$  samples usually called *frames* where  $L$  is some predetermined number (corresponding to 10-30 ms of speech). In CELP, a linear prediction (LP) filter is computed  
20                   and transmitted every frame. The  $L$ -sample frame is then divided into smaller blocks called *subframes* of size  $N$  samples, where  $L=kN$  and  $k$  is the number of subframes in a frame ( $N$  usually corresponds to 4-10 ms of speech). An excitation signal is determined in each subframe, which usually consists of two components: one from the past excitation (also  
25                   called pitch contribution or adaptive codebook) and the other from an innovation codebook (also called fixed codebook). This excitation signal

is transmitted and used at the decoder as the input of the LP synthesis filter in order to obtain the synthesized speech.

An innovation codebook in the CELP context, is an indexed set of  $N$ -sample-long sequences which will be referred to as  $N$ -dimensional codevectors. Each codebook sequence is indexed by an integer  $k$  ranging from 1 to  $M$  where  $M$  represents the size of the codebook often expressed as a number of bits  $b$ , where  $M=2^b$ .

To synthesize speech according to the CELP technique, each block of  $N$  samples is synthesized by filtering an appropriate codevector from a codebook through time varying filters modeling the spectral characteristics of the speech signal. At the encoder end, the synthetic output is computed for all, or a subset, of the codevectors from the codebook (codebook search). The retained codevector is the one producing the synthetic output closest to the original speech signal according to a perceptually weighted distortion measure. This perceptual weighting is performed using a so-called perceptual weighting filter, which is usually derived from the LP filter.

The CELP model has been very successful in encoding telephone band sound signals, and several CELP-based standards exist in a wide range of applications, especially in digital cellular applications. In the telephone band, the sound signal is band-limited to 200-3400 Hz and sampled at 8000 samples/sec. In wideband speech/audio applications, the sound signal is band-limited to 50-7000 Hz and sampled at 16000 samples/sec.

Some difficulties arise when applying the telephone-band optimized CELP model to wideband signals, and additional features need to be added to the model in order to obtain high quality wideband signals. Wideband signals exhibit a much wider dynamic range compared to telephone-band signals, which results in precision problems when a fixed-point implementation of the algorithm is required (which is essential in wireless applications). Further, the CELP model will often spend most of its encoding bits on the low-frequency region, which usually has higher energy contents, resulting in a low-pass output signal. To overcome this problem, the perceptual weighting filter has to be modified in order to suit wideband signals, and pre-emphasis techniques which boost the high frequency regions become important to reduce the dynamic range, yielding a simpler fixed-point implementation, and to ensure a better encoding of the higher frequency contents of the signal. Further, the pitch contents in the spectrum of voiced segments in wideband signals do not extend over the whole spectrum range, and the amount of voicing shows more variation compared to narrow-band signals. Therefore, in case of wideband signals, existing pitch search structures are not very efficient. Thus, it is important to improve the closed-loop pitch analysis to better accommodate the variations in the voicing level.

20

## OBJECTS OF THE INVENTION

25

An object of the present invention is therefore to provide a method and device for efficiently encoding wideband (7000 Hz) sound signals using

CELP-type encoding techniques, using improved pitch analysis in order to obtain high a quality reconstructed sound signal.

## SUMMARY OF THE INVENTION

5

More specifically, in accordance with the present invention, there is provided a method for selecting an optimal set of pitch codebook parameters associated to a signal path, from at least two signal paths, having the lowest calculated pitch prediction error. The pitch prediction error  
10 is calculated in response to a pitch codevector from a pitch codebook search device. In at least one of the two signal paths, the pitch prediction error is filtered before supplying the pitch codevector for calculation of said pitch prediction error of said one path. Finally, the pitch prediction errors calculated in said at least two signal paths are compared, the signal path  
15 having the lowest calculated pitch prediction error is chosen, and the set of pitch codebook parameters associated to the choosen signal path are selected.

The pitch analysis device of the invention, for producing an  
20 optimal set of pitch codebook parameters, comprises:

a) at least two signal paths associated to respective sets of pitch codebook parameters, wherein:

- 25
- i) each signal path comprises a pitch prediction error calculating device for calculating a pitch prediction error of a pitch codevector from a pitch codebook search device; and
  - ii) at least one of the two paths comprises a filter for filtering the pitch codevector before supplying the pitch

codevector to the path's pitch prediction error calculating device; and

b) a selector for comparing the pitch prediction errors calculated in the signal paths, for choosing the signal path having the lowest calculated pitch prediction error, and for selecting the set of pitch codebook parameters associated to the chosen signal path.

5

The new method and device which achieve efficient modeling of the harmonic structure of the speech spectrum uses several forms of low pass filters applied to the past excitation and the one yielding higher prediction gain is selected. When subsample pitch resolution is used, the low pass filters can be incorporated into the interpolation filters used to obtain the higher pitch resolution.

10

In a preferred embodiment of the invention, each pitch prediction error calculating device of the pitch analysis device described above comprises:

15

a) a convolution unit for convolving the pitch codevector with a weighted synthesis filter impulse response signal and therefore calculating a convolved pitch codevector;

20

b) a pitch gain calculator for calculating a pitch gain in response to the convolved pitch codevector and a pitch search target vector;

c) an amplifier for multiplying the convolved pitch codevector by the pitch gain to thereby produce an amplified convolved pitch codevector;

25

and



d) a combiner circuit for combining the amplified convolved pitch codevector with the pitch search target vector to thereby produce the pitch prediction error.

5 In another preferred embodiment of the invention, the pitch gain calculator comprises a means for calculating said pitch gain  $b^{(j)}$  using the relation:

$$b^{(j)} = x^T y^{(j)} / \|y^{(j)}\|^2 \quad \text{where } j = 0, 1, 2, \dots, K, \text{ and } K \text{ corresponds to a number of signal paths,}$$

10

and where  $x$  is said pitch search target vector, and  $y^{(j)}$  is said convolved pitch codevector.

15 The present invention further relates to an encoder, having the pitch analysis device described above, for encoding a wideband input signal and comprising:

- a) a linear prediction synthesis filter calculator responsive to the wideband signal for producing linear prediction synthesis filter coefficients;
- b) a perceptual weighting filter, responsive to the wideband signal
- 20 and the linear prediction synthesis filter coefficients, for producing a perceptually weighted signal;
- c) an impulse response generator responsive to the linear prediction synthesis filter coefficients for producing a weighted synthesis filter impulse response signal;
- 25 d) a pitch search unit for producing pitch codebook parameters, comprising:

i) a pitch codebook search device responsive to the perceptually weighted signal and the linear prediction synthesis filter coefficients for producing the pitch codevector and an innovative search target vector; and

5 ii) the pitch analysis device responsive to the pitch codevector for selecting, from the sets of pitch codebook parameters, the set of pitch codebook parameters associated to the path having the lowest calculated pitch prediction error;

d) an innovative codebook search device, responsive to the weighted synthesis filter impulse response signal, and the innovative search  
10 target vector, for producing innovative codebook parameters; and

e) a signal forming device for producing an encoded wideband signal comprising the set of pitch codebook parameters associated to the path having the lowest pitch prediction error, the innovative codebook parameters, and the linear prediction synthesis filter coefficients.

15

The present invention still further relates to a cellular communication system, a cellular mobile transmitter/receiver unit, a cellular network element, and a bidirectional wireless communication sub-system comprising the above described decoder.

20

The objects, advantages and other features of the present invention will become more apparent upon reading of the following non restrictive description of a preferred embodiment thereof, given by way of example only with reference to the accompanying drawings.

25

## BRIEF DESCRIPTION OF THE DRAWINGS

In the appended drawings:

5           Figure 1 is a schematic block diagram of a preferred embodiment of wideband encoding device;

          Figure 2 is a schematic block diagram of a preferred embodiment of wideband decoding device;

10

          Figure 3 is a schematic block diagram of a preferred embodiment of pitch analysis device; and

          Figure 4 is a simplified, schematic block diagram of a cellular communication system in which the wideband encoding device of Figure 1  
15           and the wideband decoding device of Figure 2 can be used.

## 20           DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

          As well known to those of ordinary skill in the art, a cellular communication system such as 401 (see Figure 4) provides a  
25           telecommunication service over a large geographic area by dividing that large geographic area into a number C of smaller cells. The C smaller cells

are serviced by respective cellular base stations  $402_1, 402_2 \dots 402_c$  to provide each cell with radio signalling, audio and data channels.

Radio signalling channels are used to page mobile radiotelephones (mobile transmitter/receiver units) such as 403 within the limits of the coverage area (cell) of the cellular base station 402, and to place calls to other radiotelephones 403 located either inside or outside the base station's cell or to another network such as the Public Switched Telephone Network (PSTN) 404.

Once a radiotelephone 403 has successfully placed or received a call, an audio or data channel is established between this radiotelephone 403 and the cellular base station 402 corresponding to the cell in which the radiotelephone 403 is situated, and communication between the base station 402 and radiotelephone 403 is conducted over that audio or data channel. The radiotelephone 403 may also receive control or timing information over a signalling channel while a call is in progress.

If a radiotelephone 403 leaves a cell and enters another adjacent cell while a call is in progress, the radiotelephone 403 hands over the call to an available audio or data channel of the new cell base station 402. If a radiotelephone 403 leaves a cell and enters another adjacent cell while no call is in progress, the radiotelephone 403 sends a control message over the signalling channel to log into the base station 402 of the new cell. In this manner mobile communication over a wide geographical area is possible.

The cellular communication system 401 further comprises a control terminal 405 to control communication between the cellular base stations

402 and the PSTN 404, for example during a communication between a radiotelephone 403 and the PSTN 404, or between a radiotelephone 403 located in a first cell and a radiotelephone 403 situated in a second cell.

Of course, a bidirectional wireless radio communication subsystem  
5 is required to establish an audio or data channel between a base station 402 of one cell and a radiotelephone 403 located in that cell. As illustrated in very simplified form in Figure 4, such a bidirectional wireless radio communication subsystem typically comprises in the radiotelephone 403:

- a transmitter 406 including:

- 10                   - an encoder 407 for encoding the voice signal; and  
                  - a transmission circuit 408 for transmitting the encoded voice signal from the encoder 407 through an antenna such as 409;  
                  and

- a receiver 410 including:

- 15                   - a receiving circuit 411 for receiving a transmitted encoded voice signal usually through the same antenna 409; and  
                  - a decoder 412 for decoding the received encoded voice signal from the receiving circuit 411.

20           The radiotelephone further comprises other conventional radiotelephone circuits 413 to which the encoder 407 and decoder 412 are connected and for processing signals therefrom, which circuits 413 are well known to those of ordinary skill in the art and, accordingly, will not be further described in the present specification.

25

Also, such a bidirectional wireless radio communication subsystem typically comprises in the base station 402:

- a transmitter 414 including:
  - an encoder 415 for encoding the voice signal; and
  - a transmission circuit 416 for transmitting the encoded voice signal from the encoder 415 through an antenna such as 417; and
- 5       - a receiver 418 including:
  - a receiving circuit 419 for receiving a transmitted encoded voice signal through the same antenna 417 or through another antenna (not shown); and
  - a decoder 420 for decoding the received encoded voice signal
- 10       from the receiving circuit 419.

The base station 402 further comprises, typically, a base station controller 421, along with its associated database 422, for controlling communication between the control terminal 405 and the transmitter 414 and  
15       receiver 418.

As well known to those of ordinary skill in the art, voice encoding is required in order to reduce the bandwidth necessary to transmit sound signal, for example voice signal such as speech, across the bidirectional  
20       wireless radio communication subsystem, i.e., between a radiotelephone 403 and a base station 402.

LP voice encoders (such as 415 and 407) typically operating at 13 kbits/second and below such as Code-Excited Linear Prediction (CELP)  
25       encoders typically use a LP synthesis filter to model the short-term spectral envelope of the voice signal. The LP information is transmitted, typically,

every 10 or 20 ms to the decoder (such 420 and 412) and is extracted at the decoder end.

The novel techniques disclosed in the present specification may apply to different LP-based coding systems. However, a CELP-type coding system is used in the preferred embodiment for the purpose of presenting a non-limitative illustration of these techniques. In the same manner, such techniques can be used with sound signals other than voice and speech as well with other types of wideband signals.

Figure 1 shows a general block diagram of a CELP-type speech encoding device 100 modified to better accommodate wideband signals.

The sampled input speech signal 114 is divided into successive  $L$ -sample blocks called "frames". In each frame, different parameters representing the speech signal in the frame are computed, encoded, and transmitted. LP parameters representing the LP synthesis filter are usually computed once every frame. The frame is further divided into smaller blocks of  $N$  samples (blocks of length  $N$ ), in which excitation parameters (pitch and innovation) are determined. In the CELP literature, these blocks of length  $N$  are called "subframes" and the  $N$ -sample signals in the subframes are referred to as  $N$ -dimensional vectors. In this preferred embodiment, the length  $N$  corresponds to 5 ms while the length  $L$  corresponds to 20 ms, which means that a frame contains four subframes ( $N=80$  at the sampling rate of 16 kHz and 64 after down-sampling to 12.8 kHz). Various  $N$ -dimensional vectors occur in the encoding procedure. A list of the vectors which appear in Figures 1 and 2 as well as a list of transmitted parameters are given herein below:

List of the main  $N$ -dimensional vectors

- s** Wideband signal input speech vector (after down-sampling, pre-processing, and preemphasis);
- 5      **$s_w$**  Weighted speech vector;
- $s_0$**  Zero-input response of weighted synthesis filter;
- $s_p$**  Down-sampled pre-processed signal;
- Oversampled synthesized speech signal;
- 10      **$s'$**  Synthesis signal before deemphasis;
- $s_d$**  Deemphasized synthesis signal;
- $s_h$**  Synthesis signal after deemphasis and postprocessing;
- $x$**  Target vector for pitch search;
- $x'$**  Target vector for innovation search;
- 15      **$h$**  Weighted synthesis filter impulse response;
- $v_T$**  Adaptive (pitch) codebook vector at delay  $T$ ;
- $y_T$**  Filtered pitch codebook vector ( $v_T$  convolved with  **$h$** );
- $c_k$**  Innovative codevector at index  $k$  ( $k$ -th entry from the innovation codebook);
- 20      **$c_r$**  Enhanced scaled innovation codevector;
- $u$**  Excitation signal (scaled innovation and pitch codevectors);
- $u'$**  Enhanced excitation;
- $z$**  Band-pass noise sequence;
- $w'$**  White noise sequence; and
- 25      **$w$**  Scaled noise sequence.



List of transmitted parameters

- STP Short term prediction parameters (defining  $A(z)$ );  
T Pitch lag (or pitch codebook index);  
b Pitch gain (or pitch codebook gain);  
5 j Index of the low-pass filter used on the pitch codevector;  
k Codevector index (innovation codebook entry); and  
g Innovation codebook gain.

10 In this preferred embodiment, the STP parameters are transmitted once per frame and the rest of the parameters are transmitted four times per frame (every subframe).

ENCODER SIDE

15 The sampled speech signal is encoded on a block by block basis by the encoding device 100 of Figure 1 which is broken down into eleven modules numbered from 101 to 111.

20 The input speech is processed into the above mentioned  $L$ -sample blocks called frames.

Referring to Figure 1, the sampled input speech signal 114 is down-sampled in a down-sampling module 101. For example, the signal is down-sampled from 16 kHz down to 12.8 kHz, using techniques well known to  
25 those of ordinary skill in the art. Down-sampling down to another frequency can of course be envisaged. Down-sampling increases the coding efficiency, since a smaller frequency bandwidth is encoded. This also

reduces the algorithmic complexity since the number of samples in a frame is decreased. The use of down-sampling becomes significant when the bit rate is reduced below 16 kbit/s, although down-sampling is not essential above 16 kbit/s.

- 5           After down-sampling, the 320-sample frame of 20 ms is reduced to 256-sample frame (down-sampling ratio of 4/5).

          The input frame is then supplied to the optional pre-processing block 102. Pre-processing block 102 may consist of a high-pass filter with a 50 Hz  
10 cut-off frequency. High-pass filter 102 removes the unwanted sound components below 50 Hz.

          The down-sampled pre-processed signal is denoted by  $s_p(n)$ ,  $n=0, 1, 2, \dots, L-1$ , where  $L$  is the length of the frame (256 at a sampling frequency of  
15 12.8 kHz). In a preferred embodiment of the preemphasis filter 103, the signal  $s_p(n)$  is preemphasized using a filter having the following transfer function:

20           
$$P(z) = 1 - \mu z^{-1}$$

          where  $\mu$  is a preemphasis factor with a value located between 0 and 1 (a  
25 typical value is  $\mu = 0.7$ ). A higher-order filter could also be used. It should be pointed out that high-pass filter 102 and preemphasis filter 103 can be interchanged to obtain more efficient fixed-point implementations.

The function of the preemphasis filter 103 is to enhance the high frequency contents of the input signal. It also reduces the dynamic range of the input speech signal, which renders it more suitable for fixed-point implementation. Without preemphasis, LP analysis in fixed-point using single-precision arithmetic is difficult to implement.

5

Preemphasis also plays an important role in achieving a proper overall perceptual weighting of the quantization error, which contributes to improved sound quality. This will be explained in more detail herein below.

10

The output of the preemphasis filter 103 is denoted  $s(n)$ . This signal is used for performing LP analysis in calculator module 104. LP analysis is a technique well known to those of ordinary skill in the art. In this preferred embodiment, the autocorrelation approach is used. In the autocorrelation approach, the signal  $s(n)$  is first windowed using a Hamming window (having usually a length of the order of 30-40 ms). The autocorrelations are computed from the windowed signal, and Levinson-Durbin recursion is used to compute LP filter coefficients,  $a_i$ , where  $i=1, \dots, p$ , and where  $p$  is the LP order, which is typically 16 in wideband coding. The parameters  $a_i$  are the coefficients of the transfer function of the LP filter, which is given by the following relation:

15

20

$$A(z) = 1 + \sum_{i=1}^p a_i z^{-i}$$

25

LP analysis is performed in calculator module 104, which also performs the quantization and interpolation of the LP filter coefficients. The LP filter coefficients are first transformed into another equivalent domain more suitable for quantization and interpolation purposes. The line spectral pair (LSP) and immitance spectral pair (ISP) domains are two domains in which quantization and interpolation can be efficiently performed. The 16 LP filter coefficients,  $a_n$ , can be quantized in the order of 30 to 50 bits using split or multi-stage quantization, or a combination thereof. The purpose of the interpolation is to enable updating the LP filter coefficients every subframe while transmitting them once every frame, which improves the encoder performance without increasing the bit rate. Quantization and interpolation of the LP filter coefficients is believed to be otherwise well known to those of ordinary skill in the art and, accordingly, will not be further described in the present specification.

The following paragraphs will describe the rest of the coding operations performed on a subframe basis. In the following description, the filter  $A(z)$  denotes the unquantized interpolated LP filter of the subframe, and the filter  $\hat{A}(z)$  denotes the quantized interpolated LP filter of the subframe.

#### **Perceptual Weighting:**

In analysis-by-synthesis encoders, the optimum pitch and innovation parameters are searched by minimizing the mean squared error between the input speech and synthesized speech in a perceptually weighted domain. This is equivalent to minimizing the error between the weighted input speech and weighted synthesis speech.

The weighted signal  $s_w(n)$  is computed in a perceptual weighting filter 105. Traditionally, the weighted signal  $s_w(n)$  is computed by a weighting filter having a transfer function  $W(z)$  in the form:

$$5 \quad W(z) = A(z/\gamma_1) / A(z/\gamma_2) \quad \text{where} \quad 0 < \gamma_2 < \gamma_1 \leq 1$$

As well known to those of ordinary skill in the art, in prior art analysis-by-synthesis (AbS) encoders, analysis shows that the quantization error is  
 10 weighted by a transfer function  $W^{-1}(z)$ , which is the inverse of the transfer function of the perceptual weighting filter 105. This result is well described by B.S. Atal and M.R. Schroeder in "Predictive coding of speech and subjective error criteria", IEEE Transaction ASSP, vol. 27, no. 3, pp. 247-254, June 1979. Transfer function  $W^{-1}(z)$  exhibits some of the formant  
 15 structure of the input speech signal. Thus, the masking property of the human ear is exploited by shaping the quantization error so that it has more energy in the formant regions where it will be masked by the strong signal energy present in these regions. The amount of weighting is controlled by the factors  $\gamma_1$  and  $\gamma_2$ .

20

The above traditional perceptual weighting filter 105 works well with telephone band signals. However, it was found that this traditional perceptual weighting filter 105 is not suitable for efficient perceptual weighting of wideband signals. It was also found that the traditional  
 25 perceptual weighting filter 105 has inherent limitations in modelling the formant structure and the required spectral tilt concurrently. The spectral tilt is more pronounced in wideband signals due to the wide dynamic range

between low and high frequencies. The prior art has suggested to add a tilt filter into  $W(z)$  in order to control the tilt and formant weighting of the wideband input signal separately.

5 A novel solution to this problem is, in accordance with the present invention, to introduce the preemphasis filter 103 at the input, compute the LP filter  $A(z)$  based on the preemphasized speech  $s(n)$ , and use a modified filter  $W(z)$  by fixing its denominator.

10 LP analysis is performed in module 104 on the preemphasized signal  $s(n)$  to obtain the LP filter  $A(z)$ . Also, a new perceptual weighting filter 105 with fixed denominator is used. An example of transfer function for the perceptual weighting filter 104 is given by the following relation:

15 
$$W(z) = A(z/\gamma_1) / (1 - \gamma_2 z^{-1}) \quad \text{where} \quad 0 < \gamma_2 < \gamma_1 \leq 1$$

A higher order can be used at the denominator. This structure substantially decouples the formant weighting from the tilt.

20

Note that because  $A(z)$  is computed based on the preemphasized speech signal  $s(n)$ , the tilt of the filter  $1/A(z/\gamma_1)$  is less pronounced compared to the case when  $A(z)$  is computed based on the original speech. Since deemphasis is performed at the decoder end using a filter having the transfer

25 function:

$$P^{-1}(z) = 1/(1 - \mu z^{-1}),$$

the quantization error spectrum is shaped by a filter having a transfer function  $W^{-1}(z)P^{-1}(z)$ . When  $\gamma_2$  is set equal to  $\mu$ , which is typically the case, the spectrum of the quantization error is shaped by a filter whose transfer function is  $1/A(z/\gamma_1)$ , with  $A(z)$  computed based on the preemphasized speech signal. Subjective listening showed that this structure for achieving the error shaping by a combination of preemphasis and modified weighting filtering is very efficient for encoding wideband signals, in addition to the advantages of ease of fixed-point algorithmic implementation.

#### Pitch Analysis:

In order to simplify the pitch analysis, an open-loop pitch lag  $T_{OL}$  is first estimated in the open-loop pitch search module 106 using the weighted speech signal  $s_w(n)$ . Then the closed-loop pitch analysis, which is performed in closed-loop pitch search module 107 on a subframe basis, is restricted around the open-loop pitch lag  $T_{OL}$  which significantly reduces the search complexity of the LTP parameters  $T$  and  $b$  (pitch lag and pitch gain). Open-loop pitch analysis is usually performed in module 106 once every 10 ms (two subframes) using techniques well known to those of ordinary skill in the art.

The target vector  $x$  for LTP (Long Term Prediction) analysis is first computed. This is usually done by subtracting the zero-input response  $s_0$  of weighted synthesis filter  $W(z)/\hat{A}(z)$  from the weighted speech signal  $s_w(n)$ . This zero-input response  $s_0$  is calculated by a zero-input response calculator

108. More specifically, the target vector  $x$  is calculated using the following relation:

$$x = s_w - s_0$$

5

where  $x$  is the  $N$ -dimensional target vector,  $s_w$  is the weighted speech vector in the subframe, and  $s_0$  is the zero-input response of filter  $W(z)/\hat{A}(z)$  which is the output of the combined filter  $W(z)/\hat{A}(z)$  due to its initial states.

10 The zero-input response calculator 108 is responsive to the quantized interpolated LP filter  $\hat{A}(z)$  from the LP analysis, quantization and interpolation calculator 104 and to the initial states of the weighted synthesis filter  $W(z)/\hat{A}(z)$  stored in memory module 111 to calculate the zero-input response  $s_0$  (that part of the response due to the initial states as determined by setting  
15 the inputs equal to zero) of filter  $W(z)/\hat{A}(z)$ . This operation is well known to those of ordinary skill in the art and, accordingly, will not be further described.

Of course, alternative but mathematically equivalent approaches can be used to compute the target vector  $x$ .

20

A  $N$ -dimensional impulse response vector  $h$  of the weighted synthesis filter  $W(z)/\hat{A}(z)$  is computed in the impulse response generator 109 using the LP filter coefficients  $A(z)$  and  $\hat{A}(z)$  from module 104. Again, this operation is well known to those of ordinary skill in the art and, accordingly,  
25 will not be further described in the present specification.

The closed-loop pitch (or pitch codebook) parameters  $b$ ,  $T$  and  $j$  are



computed in the closed-loop pitch search module 107, which uses the target vector  $x$ , the impulse response vector  $h$  and the open-loop pitch lag  $T_{ol}$  as inputs. Traditionally, the pitch prediction has been represented by a pitch filter having the following transfer function:

5

$$1 / (1 - bz^{-T})$$

where  $b$  is the pitch gain and  $T$  is the pitch delay or lag. In this case, the pitch contribution to the excitation signal  $u(n)$  is given by  $bu(n-T)$ , where the total excitation is given by

10

$$u(n) = bu(n-T) + gc_k(n)$$

15

with  $g$  being the innovative codebook gain and  $c_k(n)$  the innovative codevector at index  $k$ .

20 This representation has limitations if the pitch lag  $T$  is shorter than the subframe length  $N$ . In another representation, the pitch contribution can be seen as a pitch codebook containing the past excitation signal. Generally, each vector in the pitch codebook is a shift-by-one version of the previous vector (discarding one sample and adding a new sample). For pitch lags  
25  $T > N$ , the pitch codebook is equivalent to the filter structure  $(1/(1-bz^T))$ , and a pitch codebook vector  $v_T(n)$  at pitch lag  $T$  is given by

$$v_T(n) = u(n-T) \quad ; \quad n=0, \dots, N-1.$$

For pitch lags  $T$  shorter than  $N$ , a vector  $v_T(n)$  is built by repeating the  
 5 available samples from the past excitation until the vector is completed (this is not equivalent to the filter structure).

In recent encoders, a higher pitch resolution is used which significantly improves the quality of voiced sound segments. This is  
 10 achieved by oversampling the past excitation signal using polyphase interpolation filters. In this case, the vector  $v_T(n)$  usually corresponds to an interpolated version of the past excitation, with pitch lag  $T$  being a non-integer delay (e.g. 50.25).

15 The pitch search consists of finding the best pitch lag  $T$  and gain  $b$  that minimize the mean squared weighted error  $E$  between the target vector  $x$  and the scaled filtered past excitation. Error  $E$  being expressed as:

$$20 \quad E = \|x - by_T\|^2$$

where  $y_T$  is the filtered pitch codebook vector at pitch lag  $T$ :

$$25 \quad y_T(n) = v_T(n) * h(n) = \sum_{i=0}^n v_T(i)h(n-i) \quad , \quad n=0, \dots, N-1.$$

It can be shown that the error  $E$  is minimized by maximizing the search criterion

5

$$C = \frac{x^t y_T}{\sqrt{y_T^t y_T}}$$

10 where  $t$  denotes vector transpose.

In the preferred embodiment of the present invention, a 1/3 subsample pitch resolution is used, and the pitch (pitch codebook) search is composed of three stages.

15

In the first stage, an open-loop pitch lag  $T_{OL}$  is estimated in open-loop pitch search module 106 in response to the weighted speech signal  $s_w(n)$ . As indicated in the foregoing description, this open-loop pitch analysis is usually performed once every 10 ms (two subframes) using techniques well  
20 known to those of ordinary skill in the art.

25

In the second stage, the search criterion  $C$  is searched in the closed-loop pitch search module 107 for integer pitch lags around the estimated open-loop pitch lag  $T_{OL}$  (usually  $\pm 5$ ), which significantly simplifies the search procedure. A simple procedure is used for updating the filtered codevector  $y_T$  without the need to compute the convolution for every pitch lag.

Once an optimum integer pitch lag is found in the second stage, a third stage of the search (module 107) tests the fractions around that optimum integer pitch lag.

When the pitch predictor is represented by a filter of the form  $1/(1-bz^{-T})$ , which is a valid assumption for pitch lags  $T > N$ , the spectrum of the pitch filter exhibits a harmonic structure over the entire frequency range, with a harmonic frequency related to  $1/T$ . In case of wideband signals, this structure is not very efficient since the harmonic structure in wideband signals does not cover the entire extended spectrum. The harmonic structure exists only up to a certain frequency, depending on the speech segment. Thus, in order to achieve efficient representation of the pitch contribution in voiced segments of wideband speech, the pitch prediction filter needs to have the flexibility of varying the amount of periodicity over the wideband spectrum.

A new method which achieves efficient modeling of the harmonic structure of the speech spectrum of wideband signals is disclosed in the present specification, whereby several forms of low pass filters are applied to the past excitation and the low pass filter with higher prediction gain is selected.

When subsample pitch resolution is used, the low pass filters can be incorporated into the interpolation filters used to obtain the higher pitch resolution. In this case, the third stage of the pitch search, in which the fractions around the chosen integer pitch lag are tested, is repeated for the several interpolation filters having different low-pass characteristics and the fraction and filter index which maximize the search criterion  $C$  are selected.

A simpler approach is to complete the search in the three stages described above to determine the optimum fractional pitch lag using only one interpolation filter with a certain frequency response, and select the optimum low-pass filter shape at the end by applying the different predetermined low-pass filters to the chosen pitch codebook vector  $v_T$  and select the low-pass filter which minimizes the pitch prediction error. This approach is discussed in detail below.

Figure 3 illustrates a schematic block diagram of a preferred embodiment of the proposed approach.

10

In memory module 303, the past excitation signal  $u(n)$ ,  $n < 0$ , is stored. The pitch codebook search module 301 is responsive to the target vector  $x$ , to the open-loop pitch lag  $T_{OL}$  and to the past excitation signal  $u(n)$ ,  $n < 0$ , from memory module 303 to conduct a pitch codebook (pitch codebook) search minimizing the above-defined search criterion  $C$ . From the result of the search conducted in module 301, module 302 generates the optimum pitch codebook vector  $v_T$ . Note that since a sub-sample pitch resolution is used (fractional pitch), the past excitation signal  $u(n)$ ,  $n < 0$ , is interpolated and the pitch codebook vector  $v_T$  corresponds to the interpolated past excitation signal. In this preferred embodiment, the interpolation filter (in module 301, but not shown) has a low-pass filter characteristic removing the frequency contents above 7000 Hz.

15

20

In a preferred embodiment,  $K$  filter characteristics are used; these filter characteristics could be low-pass or band-pass filter characteristics. Once the optimum codevector  $v_T$  is determined and supplied by the pitch codevector generator 302,  $K$  filtered versions of  $v_T$  are computed

25

respectively using  $K$  different frequency shaping filters such as 305<sup>(j)</sup>, where  $j=1, 2, \dots, K$ . These filtered versions are denoted  $v_f^{(j)}$ , where  $j=1, 2, \dots, K$ . The different vectors  $v_f^{(j)}$  are convolved in respective modules 304<sup>(j)</sup>, where  $j=0, 1, 2, \dots, K$ , with the impulse response  $h$  to obtain the vectors  $y^{(j)}$ , where  $j=0, 1, 2, \dots, K$ . To calculate the mean squared pitch prediction error for each vector  $y^{(j)}$ , the value  $y^{(j)}$  is multiplied by the gain  $b$  by means of a corresponding amplifier 307<sup>(j)</sup> and the value  $by^{(j)}$  is subtracted from the target vector  $x$  by means of a corresponding subtractor 308<sup>(j)</sup>. Selector 309 selects the frequency shaping filter 305<sup>(j)</sup> which minimizes the mean squared pitch prediction error

$$e^{(j)} = \|x - b^{(j)} y^{(j)}\|^2, \quad j=1, 2, \dots, K$$

To calculate the mean squared pitch prediction error  $e^{(j)}$  for each value of  $y^{(j)}$ , the value  $y^{(j)}$  is multiplied by the gain  $b$  by means of a corresponding amplifier 307<sup>(j)</sup> and the value  $b^{(j)} y^{(j)}$  is subtracted from the target vector  $x$  by means of subtractors 308<sup>(j)</sup>. Each gain  $b^{(j)}$  is calculated in a corresponding gain calculator 306<sup>(j)</sup> in association with the frequency shaping filter at index  $j$ , using the following relationship:

$$b^{(j)} = x^T y^{(j)} / \|y^{(j)}\|^2$$

In selector 309, the parameters  $b$ ,  $T$ , and  $j$  are chosen based on  $v_T$  or  $v_f^{(j)}$  which minimizes the mean squared pitch prediction error  $e$ .

Referring back to Figure 1, the pitch codebook index  $T$  is encoded and transmitted to multiplexer 112. The pitch gain  $b$  is quantized and transmitted to multiplexer 112. With this new approach, extra information is needed to encode the index  $j$  of the selected frequency shaping filter in multiplexer 112. For example, if three filters are used ( $j=0, 1, 2, 3$ ), then two bits are needed to represent this information. The filter index information  $j$  can also be encoded jointly with the pitch gain  $b$ .

#### Innovative codebook search:

Once the pitch, or LTP (Long Term Prediction) parameters  $b$ ,  $T$ , and  $j$  are determined, the next step is to search for the optimum innovative excitation by means of search module 110 of Figure 1. First, the target vector  $x$  is updated by subtracting the LTP contribution:

$$x' = x - by_T$$

where  $b$  is the pitch gain and  $y_T$  is the filtered pitch codebook vector (the past excitation at delay  $T$  filtered with the selected low pass filter and convolved with the impulse response  $h$  as described with reference to Figure 3).

The search procedure in CELP is performed by finding the optimum excitation codevector  $c_k$  and gain  $g$  which minimize the mean-squared error between the target vector and the scaled filtered codevector

$$E = \| x' - gHc_k \|^2$$

where  $H$  is a lower triangular convolution matrix derived from the impulse  
5 response vector  $h$ .

In the preferred embodiment of the present invention, the innovative  
codebook search is performed in module 110 by means of an algebraic  
codebook as described in US patents Nos: 5,444,816 (Adoul et al.) issued  
10 on August 22, 1995; 5,699,482 granted to Adoul et al., on December 17,  
1997; 5,754,976 granted to Adoul et al., on May 19, 1998; and 5,701,392  
(Adoul et al.) dated December 23, 1997.

Once the optimum excitation codevector  $c_k$  and its gain  $g$  are chosen  
15 by module 110, the codebook index  $k$  and gain  $g$  are encoded and  
transmitted to multiplexer 112.

Referring to Figure 1, the parameters  $b$ ,  $T$ ,  $j$ ,  $\hat{A}(z)$ ,  $k$  and  $g$  are  
multiplexed through the multiplexer 112 before being transmitted through a  
20 communication channel.

#### Memory update:

25 In memory module 111 (Figure 1), the states of the weighted  
synthesis filter  $W(z)/\hat{A}(z)$  are updated by filtering the excitation signal



$u = gc_k + bv_T$  through the weighted synthesis filter. After this filtering, the states of the filter are memorized and used in the next subframe as initial states for computing the zero-input response in calculator module 108.

5 As in the case of the target vector  $x$ , other alternative but mathematically equivalent approaches well known to those of ordinary skill in the art can be used to update the filter states.

### DECODER SIDE

10

The speech decoding device 200 of Figure 2 illustrates the various steps carried out between the digital input 222 (input stream to the demultiplexer 217) and the output sampled speech 223 (output of the adder 221).

15

Demultiplexer 217 extracts the synthesis model parameters from the binary information received from a digital input channel. From each received binary frame, the extracted parameters are:

- 20
- the short-term prediction parameters (STP)  $\hat{A}(z)$  (once per frame);
  - the long-term prediction (LTP) parameters  $T$ ,  $b$ , and  $j$  (for each subframe); and
- 25
- the innovation codebook index  $k$  and gain  $g$  (for each subframe).

The current speech signal is synthesized based on these parameters as will be explained hereinbelow.

5           The innovative codebook 218 is responsive to the index  $k$  to produce the innovation codevector  $c_k$ , which is scaled by the decoded gain factor  $g$  through an amplifier 224. In the preferred embodiment, an innovative codebook 218 as described in the above mentioned US patent numbers 5,444,816; 5,699,482; 5,754,976; and 5,701,392 is used to represent the innovative codevector  $c_k$ .

10           The generated scaled codevector  $gc_k$  at the output of the amplifier 224 is processed through a innovation filter 205.

#### **Periodicity enhancement:**

15           The generated scaled codevector at the output of the amplifier 224 is processed through a frequency-dependent pitch enhancer 205.

20           Enhancing the periodicity of the excitation signal  $u$  improves the quality in case of voiced segments. This was done in the past by filtering the innovation vector from the innovative codebook (fixed codebook) 218 through a filter in the form  $1/(1-\varepsilon bz^{-T})$  where  $\varepsilon$  is a factor below 0.5 which controls the amount of introduced periodicity. This approach is less efficient in case of wideband signals since it introduces periodicity over the entire spectrum. A new alternative approach, which is part of the present invention, is disclosed whereby periodicity enhancement is  
25           achieved by filtering the innovative codevector  $c_k$  from the innovative (fixed) codebook through an innovation filter 205 ( $F(z)$ ) whose frequency

response emphasizes the higher frequencies more than lower frequencies. The coefficients of  $F(z)$  are related to the amount of periodicity in the excitation signal  $u$ .

5 Many methods known to those skilled in the art are available for obtaining valid periodicity coefficients. For example, the value of gain  $b$  provides an indication of periodicity. That is, if gain  $b$  is close to 1, the periodicity of the excitation signal  $u$  is high, and if gain  $b$  is less than 0.5, then periodicity is low.

10 Another efficient way to derive the filter  $F(z)$  coefficients used in a preferred embodiment, is to relate them to the amount of pitch contribution in the total excitation signal  $u$ . This results in a frequency response depending on the subframe periodicity, where higher frequencies are more strongly emphasized (stronger overall slope) for  
15 higher pitch gains. Innovation filter 205 has the effect of lowering the energy of the innovative codevector  $c_k$  at low frequencies when the excitation signal  $u$  is more periodic, which enhances the periodicity of the excitation signal  $u$  at lower frequencies more than higher frequencies. Suggested forms for innovation filter 205 are

20

$$(1) F(z)=1-\sigma z^{-1}, \quad \text{or} \quad (2) F(z)=-\alpha z+1-\alpha z^{-1}$$

25 where  $\sigma$  or  $\alpha$  are periodicity factors derived from the level of periodicity of the excitation signal  $u$ .

The second three-term form of  $F(z)$  is used in a preferred embodiment. The periodicity factor  $\alpha$  is computed in the voicing factor generator 204. Several methods can be used to derive the periodicity factor  $\alpha$  based on the periodicity of the excitation signal  $u$ . Two methods are presented below.

5

#### Method 1:

The ratio of pitch contribution to the total excitation signal  $u$  is first computed in voicing factor generator 204 by

10

$$R_p = \frac{b^2 v_T^T v_T}{u^T u} = \frac{b^2 \sum_{n=0}^{N-1} v_T^2(n)}{\sum_{n=0}^{N-1} u^2(n)}$$

15

where  $v_T$  is the pitch codebook vector,  $b$  is the pitch gain, and  $u$  is the excitation signal  $u$  given at the output of the adder 219 by

$$20 \quad u = gc_k + bv_T$$

Note that the term  $bv_T$  has its source in the pitch codebook (pitch codebook) 201 in response to the pitch lag  $T$  and the past value of  $u$  stored in memory 203. The pitch codevector  $v_T$  from the pitch codebook 201 is then processed through a low-pass filter 202 whose cut-off frequency is adjusted by means of the index  $j$  from the demultiplexer 217. The resulting codevector  $v_T$  is then multiplied by the gain  $b$  from the

25

demultiplexer 217 through an amplifier 226 to obtain the signal  $bv_T$ .

The factor  $\alpha$  is calculated in voicing factor generator 204 by

$$\alpha = qR_p \quad \text{bounded by} \quad \alpha < q$$

5

where  $q$  is a factor which controls the amount of enhancement ( $q$  is set to 0.25 in this preferred embodiment).

#### Method 2:

10

Another method used in a preferred embodiment of the invention for calculating periodicity factor  $\alpha$  is discussed below.

15 by

First, a voicing factor  $r_v$  is computed in voicing factor generator 204

$$r_v = (E_v - E_c) / (E_v + E_c)$$

20 where  $E_v$  is the energy of the scaled pitch codevector  $bv_T$  and  $E_c$  is the energy of the scaled innovative codevector  $gc_k$ . That is

$$E_v = b^2 v_T^t v_T = b^2 \sum_{n=0}^{N-1} v_T^2(n)$$

25 and

$$E_c = g^2 c_k^t c_k = g^2 \sum_{n=0}^{N-1} c_k^2(n)$$

Note that the value of  $r_v$  lies between -1 and 1 (1 corresponds to purely voiced signals and -1 corresponds to purely unvoiced signals).

In this preferred embodiment, the factor  $\alpha$  is then computed in voicing factor generator 204 by

5

$$\alpha = 0.125 (1 + r_v)$$

10 which corresponds to a value of 0 for purely unvoiced signals and 0.25 for purely voiced signals.

In the first, two-term form of  $F(z)$ , the periodicity factor  $\sigma$  can be approximated by using  $\sigma = 2\alpha$  in methods 1 and 2 above. In such a case, the periodicity factor  $\sigma$  is calculated as follows in method 1 above:

15

$$\sigma = 2qR_p \quad \text{bounded by} \quad \sigma < 2q.$$

In method 2, the periodicity factor  $\sigma$  is calculated as follows:

$$\sigma = 0.25 (1 + r_v).$$

20

The enhanced signal  $c_r$  is therefore computed by filtering the scaled innovative codevector  $gc_k$  through the innovation filter 205 ( $F(z)$ ).

The enhanced excitation signal  $u'$  is computed by the adder 220

25

as:

$$u' = c_f + bv_T$$

Note that this process is not performed at the encoder 100. Thus, it is essential to update the content of the pitch codebook 201 using the excitation signal  $u$  without enhancement to keep synchronism between the encoder 100 and decoder 200. Therefore, the excitation signal  $u$  is used to update the memory 203 of the pitch codebook 201 and the enhanced excitation signal  $u'$  is used at the input of the LP synthesis filter 206.

### Synthesis and deemphasis

The synthesized signal  $s'$  is computed by filtering the enhanced excitation signal  $u'$  through the LP synthesis filter 206 which has the form  $1/\hat{A}(z)$ , where  $\hat{A}(z)$  is the interpolated LP filter in the current subframe. As can be seen in Figure 2, the quantized LP coefficients  $\hat{A}(z)$  on line 225 from demultiplexer 217 are supplied to the LP synthesis filter 206 to adjust the parameters of the LP synthesis filter 206 accordingly. The deemphasis filter 207 is the inverse of the preemphasis filter 103 of Figure 1. The transfer function of the deemphasis filter 207 is given by

$$D(z) = 1 / (1 - \mu z^{-1})$$

where  $\mu$  is a preemphasis factor with a value located between 0 and 1 (a typical value is  $\mu = 0.7$ ). A higher-order filter could also be used.

The vector  $\mathbf{s}'$  is filtered through the deemphasis filter  $D(z)$  (module 207) to obtain the vector  $\mathbf{s}_o$ , which is passed through the high-pass filter 208 to remove the unwanted frequencies below 50 Hz and further obtain  $\mathbf{s}_h$ .

### Oversampling and high-frequency regeneration

The over-sampling module 209 conducts the inverse process of the down-sampling module 101 of Figure 1. In this preferred embodiment, oversampling converts from the 12.8 kHz sampling rate to the original 16 kHz sampling rate, using techniques well known to those of ordinary skill in the art. The oversampled synthesis signal is denoted  $\hat{\mathbf{S}}$ . Signal  $\hat{\mathbf{S}}$  is also referred to as the synthesized wideband intermediate signal.

The oversampled synthesis  $\hat{\mathbf{S}}$  signal does not contain the higher frequency components which were lost by the downsampling process (module 101 of Figure 1) at the encoder 100. This gives a low-pass perception to the synthesized speech signal. To restore the full band of the original signal, a high frequency generation procedure is disclosed. This procedure is performed in modules 210 to 216, and adder 221, and requires input from voicing factor generator 204 (Figure 2).

In this new approach, the high frequency contents are generated by filling the upper part of the spectrum with a white noise properly scaled in the excitation domain, then converted to the speech domain, preferably by



shaping it with the same LP synthesis filter used for synthesizing the down-sampled signal  $\hat{S}$ .

The high frequency generation procedure in accordance with the present invention is described hereinbelow.

5

The random noise generator 213 generates a white noise sequence  $w'$  with a flat spectrum over the entire frequency bandwidth, using techniques well known to those of ordinary skill in the art. The generated sequence is of length  $N'$  which is the subframe length in the original domain.

10 Note that  $N$  is the subframe length in the down-sampled domain. In this preferred embodiment,  $N=64$  and  $N'=80$  which correspond to 5 ms.

The white noise sequence is properly scaled in the gain adjusting module 214. Gain adjustment comprises the following steps. First, the energy of the generated noise sequence  $w'$  is set equal to the energy of the enhanced excitation signal  $u'$  computed by an energy computing module 210, and the resulting scaled noise sequence is given by

20

$$w(n) = w'(n) \sqrt{\frac{\sum_{n=0}^{N-1} u'^2(n)}{\sum_{n=0}^{N'-1} w'^2(n)}}, \quad n=0, \dots, N'-1.$$

25 The second step in the gain scaling is to take into account the high frequency contents of the synthesized signal at the output of the voicing factor generator 204 so as to reduce the energy of the generated noise in

case of voiced segments (where less energy is present at high frequencies compared to unvoiced segments). In this preferred embodiment, measuring the high frequency contents is implemented by measuring the tilt of the synthesis signal through a spectral tilt calculator 212 and reducing the energy accordingly. Other measurements such as zero crossing

5 measurements can equally be used. When the tilt is very strong, which corresponds to voiced segments, the noise energy is further reduced. The tilt factor is computed in module 212 as the first correlation coefficient of the synthesis signal  $s_h$  and it is given by:

10

$$tilt = \frac{\sum_{n=1}^{N-1} s_h(n) s_h(n-1)}{\sum_{n=0}^{N-1} s_h^2(n)}, \quad \text{conditioned by } tilt \geq 0 \text{ and } tilt \geq r_v$$

15

where voicing factor  $r_v$  is given by

20

$$r_v = (E_v - E_c) / (E_v + E_c)$$

25

where  $E_v$  is the energy of the scaled pitch codevector  $bv_T$  and  $E_c$  is the energy of the scaled innovative codevector  $gc_k$ , as described earlier. Voicing factor  $r_v$  is most often less than  $tilt$  but this condition was introduced as a

precaution against high frequency tones where the tilt value is negative and the value of  $r_v$  is high. Therefore, this condition reduces the noise energy for such tonal signals.

5 The tilt value is 0 in case of flat spectrum and 1 in case of strongly voiced signals, and it is negative in case of unvoiced signals where more energy is present at high frequencies.

10 Different methods can be used to derive the scaling factor  $g_t$  from the amount of high frequency contents. In this invention, two methods are given based on the tilt of signal described above.

#### Method 1:

15 The scaling factor  $g_t$  is derived from the tilt by

$$g_t = 1 - \text{tilt} \quad \text{bounded by} \quad 0.2 \leq g_t \leq 1.0$$

20 For strongly voiced signal where the tilt approaches 1,  $g_t$  is 0.2 and for strongly unvoiced signals  $g_t$  becomes 1.0.

#### Method 2:

25 The tilt factor  $g_t$  is first restricted to be larger or equal to zero, then the scaling factor is derived from the tilt by

$$g_t = 10^{-0.8 \text{tilt}}$$

The scaled noise sequence  $w_g$  produced in gain adjusting module 214 is therefore given by:

5

$$w_g = g_t w.$$

When the tilt is close to zero, the scaling factor  $g_t$  is close to 1, which does not result in energy reduction. When the tilt value is 1, the scaling factor  $g_t$  results in a reduction of 12 dB in the energy of the generated noise.

10

Once the noise is properly scaled ( $w_g$ ), it is brought into the speech domain using the spectral shaper 215. In the preferred embodiment, this is achieved by filtering the noise  $w_g$  through a bandwidth expanded version of the same LP synthesis filter used in the down-sampled domain ( $1/\hat{A}(z/0.8)$ ). The corresponding bandwidth expanded LP filter coefficients are calculated in spectral shaper 215.

15

The filtered scaled noise sequence  $w_f$  is then band-pass filtered to the required frequency range to be restored using the band-pass filter 216. In the preferred embodiment, the band-pass filter 216 restricts the noise sequence to the frequency range 5.6-7.2 kHz. The resulting band-pass filtered noise sequence  $z$  is added in adder 221 to the oversampled synthesized speech signal  $\hat{s}$  to obtain the final reconstructed sound signal  $s_{out}$  on the output 223.

20

25

Although the present invention has been described hereinabove by way of a preferred embodiment thereof, this embodiment can be modified at will, within the scope of the appended claims, without departing from the spirit and nature of the subject invention. Even though the preferred embodiment discusses the use of wideband speech signals, it will be  
5 obvious to those skilled in the art that the subject invention is also directed to other embodiments using wideband signals in general and that it is not necessarily limited to speech applications.

**WHAT IS CLAIMED IS:**

1. A pitch analysis device for producing an optimal set of pitch codebook parameters, comprising:
  - a) at least two signal paths associated to respective sets of pitch codebook parameters, wherein:
    - i) each signal path comprises a pitch prediction error calculating device for calculating a pitch prediction error of a pitch codevector from a pitch codebook search device; and
    - ii) at least one of said two paths comprises a filter for filtering the pitch codevector before supplying said pitch codevector to the pitch prediction error calculating device of said one path; and
  - b) a selector for comparing the pitch prediction errors calculated in said at least two signal paths, for choosing the signal path having the lowest calculated pitch prediction error, and for selecting the set of pitch codebook parameters associated to the chosen signal path.
2. A pitch analysis device as defined in claim 1, wherein one of said at least two paths comprises no filter for filtering the pitch codevector before supplying said pitch codevector to the pitch prediction error calculating device.
3. A pitch analysis device as defined in claim 1, wherein said signal paths comprises a plurality of signal paths each provided with a filter for filtering the pitch codevector before supplying said pitch codevector to the pitch prediction error calculating device of the same path.

4. A pitch analysis device as defined in claim 3, wherein the filters of said plurality of paths are selected from the group consisting of low-pass and band-pass filters, and wherein said filters have different frequency responses.

5. A pitch analysis device as defined in claim 1, wherein each pitch prediction error calculating device comprises:

a) a convolution unit for convolving the pitch codevector with a weighted synthesis filter impulse response signal and therefore calculating a convolved pitch codevector;

b) a pitch gain calculator for calculating a pitch gain in response to the convolved pitch codevector and a pitch search target vector;

c) an amplifier for multiplying the convolved pitch codevector by the pitch gain to thereby produce an amplified convolved pitch codevector; and

d) a combiner circuit for combining the amplified convolved pitch codevector with the pitch search target vector to thereby produce the pitch prediction error.

6. A pitch analysis device as defined in claim 5, wherein said pitch gain calculator comprises a means for calculating said pitch gain  $b^{(j)}$  using the relation:

$$b^{(j)} = x^t y^{(j)} / \|y^{(j)}\|^2 \quad \text{where } j = 0, 1, 2, \dots, K, \text{ and } K \text{ corresponds to a number of signal paths,}$$

and where  $x$  is said pitch search target vector, and  $y^{(j)}$  is said convolved pitch codevector.

7. A pitch analysis device as defined in claim 1, wherein said pitch prediction error calculating device of each signal path comprises means for calculating an energy of the corresponding pitch prediction error, and wherein said selector comprises means for comparing the energies of said pitch prediction errors of the different signal paths and for choosing as the  
5 signal path having the lowest calculated pitch prediction error the signal path having the lowest calculated energy of the pitch prediction error.
8. A pitch analysis device as defined in claim 5, wherein:
- 10 a) each of said filters of the plurality of signal paths is identified by a filter index;
- b) said pitch codevector is identified by a pitch codebook index; and
- c) said pitch codebook parameters comprise the filter index, the pitch codebook index and the pitch gain.
- 15 9. A pitch analysis device as defined in claim 1, wherein said filter is integrated in an interpolation filter of said pitch codebook search device, said interpolation filter being used to produce a sub-sample version of said pitch codevector.
- 20 10. A pitch analysis method for producing an optimal set of pitch codebook parameters, comprising:
- a) in at least two signal paths associated to respective sets of pitch codebook parameters, calculating, for each signal path, a pitch prediction error of a pitch codevector from a pitch codebook search device;
- 25 b) in at least one of said two signal paths, filtering the pitch codevector before supplying said pitch codevector for calculation of said pitch prediction error of said one path; and



c) comparing the pitch prediction errors calculated in said at least two signal paths, choosing the signal path having the lowest calculated pitch prediction error, and selecting the set of pitch codebook parameters associated to the choosen signal path.

- 5      11. A pitch analysis method as defined in claim 10, wherein, in one of said at least two paths, no filtering of the pitch codevector is performed before supplying said pitch codevector to the pitch prediction error calculating device.
- 10     12. A pitch analysis method as defined in claim 10, wherein said signal paths comprises a plurality of signal paths and wherein filtering the pitch codevector is performed in each of said plurality of signal paths before supplying said pitch codevector to the pitch prediction error calculating device of the same path.
- 15     13. A pitch analysis method as defined in claim 12, further comprising selecting the filters of said plurality of paths from the group consisting of low-pass and band-pass filters, and wherein said filters have different frequency responses.
- 20     14. A pitch analysis method as defined in claim 10, wherein calculating a pitch prediction error in each signal path comprises:
- 25       a) convolving the pitch codevector with a weighted synthesis filter impulse response signal and therefore calculating a convolved pitch codevector;
- b) calculating a pitch gain in response to the convolved pitch codevector and a pitch search target vector;

c) multiplying the convolved pitch codevector by the pitch gain to thereby produce an amplified convolved pitch codevector; and

d) combining the amplified convolved pitch codevector with the pitch search target vector to thereby produce the pitch prediction error.

- 5      15. A pitch analysis method as defined in claim 14, wherein said pitch gain calculation comprises calculating said pitch gain  $b^{(j)}$  using the relation:

$$b^{(j)} = x^T y^{(j)} / \|y^{(j)}\|^2 \quad \text{where } j = 0, 1, 2, \dots, K, \text{ and } K \text{ corresponds to a}$$

number of signal paths,

10

and where  $x$  is said pitch search target vector, and  $y^{(j)}$  is said convolved pitch codevector.

- 15      16. A pitch analysis method as defined in claim 10, wherein calculating said pitch prediction error, in each signal path, comprises calculating an energy of the corresponding pitch prediction error, and wherein comparing the pitch prediction error comprises comparing the energies of said pitch prediction errors of the different signal paths and choosing as the signal path having the lowest calculated pitch prediction error the signal path having the lowest
- 20      calculated energy of the pitch prediction error.

17. A pitch analysis method as defined in claim 14, wherein:

a) identifying each of said filters of the plurality of signal paths by a filter index;

- 25      b) identifying said pitch codevector by a pitch codebook index; and

c) said pitch codebook parameters comprise the filter index, the pitch codebook index and the pitch gain.

18. A pitch analysis method as defined in claim 10, wherein said filtering the pitch codevector is integrated in an interpolation filter of said pitch codebook search device, said interpolation filter being used to produce a sub-sample version of said pitch codevector.

5 19. An encoder having a pitch analysis device as in claim 1 for encoding a wideband input signal, said encoder comprising:

a) a linear prediction synthesis filter calculator responsive to the wideband signal for producing linear prediction synthesis filter coefficients;

10 b) a perceptual weighting filter, responsive to the wideband signal and the linear prediction synthesis filter coefficients, for producing a perceptually weighted signal;

c) an impulse response generator responsive to said linear prediction synthesis filter coefficients for producing a weighted synthesis filter impulse response signal;

15 d) a pitch search unit for producing pitch codebook parameters, said pitch search unit comprising:

20 i) said pitch codebook search device responsive to the perceptually weighted signal and the linear prediction synthesis filter coefficients for producing the pitch codevector and an innovative search target vector; and

ii) said pitch analysis device responsive to the pitch codevector for selecting, from said sets of pitch codebook parameters, the set of pitch codebook parameters associated to the path having the lowest calculated pitch prediction error;

25 d) an innovative codebook search device, responsive to the weighted synthesis filter impulse response signal, and the innovative search target vector, for producing innovative codebook parameters; and

e) a signal forming device for producing an encoded wideband signal comprising the set of pitch codebook parameters associated to the path having the lowest pitch prediction error, said innovative codebook parameters, and said linear prediction synthesis filter coefficients.

5      20. An encoder as defined in claim 19, wherein one of said at least two paths comprises no filter for filtering the pitch codevector before supplying said pitch codevector to the pitch prediction error calculating device.

10      21. An encoder as defined in claim 19, wherein said signal paths comprises a plurality of signal paths each provided with a filter for filtering the pitch codevector before supplying said pitch codevector to the pitch prediction error calculating device of the same path.

15      22. An encoder as defined in claim 21, wherein the filters of said plurality of paths are selected from the group consisting of low-pass and band-pass filters, and wherein said filters have different frequency responses.

20      23. An encoder as defined in claim 19, wherein each pitch prediction error calculating device comprises:

20          a) a convolution unit for convolving the pitch codevector with the weighted synthesis filter impulse response signal and therefore calculating a convolved pitch codevector;

        b) a pitch gain calculator for calculating a pitch gain in response to the convolved pitch codevector and the pitch search target vector;

25          c) an amplifier for multiplying the convolved pitch codevector by the pitch gain to thereby produce an amplified convolved pitch codevector; and

d) a combiner circuit for combining the amplified convolved pitch codevector with the pitch search target vector to thereby produce the pitch prediction error.

24. An encoder as defined in claim 23, wherein said pitch gain calculator  
5 comprises a means for calculating said pitch gain  $b^{(j)}$  using the relation:

$$b^{(j)} = \mathbf{x}^T \mathbf{y}^{(j)} / \|\mathbf{y}^{(j)}\|^2 \quad \text{where } j = 0, 1, 2, \dots, K, \text{ and } K \text{ corresponds to a number of signal paths,}$$

- 10 and where  $\mathbf{x}$  is said pitch search target vector, and  $\mathbf{y}^{(j)}$  is said convolved pitch codevector.

25. An encoder as defined in claim 19, wherein said pitch prediction error calculating device of each signal path comprises means for calculating an  
15 energy of the corresponding pitch prediction error, and wherein said selector comprises means for comparing the energies of said pitch prediction errors of the different signal paths and for choosing as the signal path having the lowest calculated pitch prediction error the signal path having the lowest calculated energy of the pitch prediction error.

20

26. An encoder as defined in claim 23, wherein:

a) each of said filters of the plurality of signal paths is identified by a filter index;

b) said pitch codevector is identified by a pitch codebook index; and

- 25 c) said pitch codebook parameters comprise the filter index, the pitch codebook index and the pitch gain.

27. A encoder as defined in claim 19, wherein said filter is integrated in an interpolation filter of said pitch codebook search device, said interpolation filter being used to produce a sub-sample version of said pitch codevector.

28. A cellular communication system for servicing a large geographical area divided into a plurality of cells, comprising:

5

- a) mobile transmitter/receiver units;
- b) cellular base stations respectively situated in said cells;
- c) a control terminal for controlling communication between the cellular base stations;

10

d) a bidirectional wireless communication sub-system between each mobile unit situated in one cell and the cellular base station of said one cell, said bidirectional wireless communication sub-system comprising, in both the mobile unit and the cellular base station:

15

i) a transmitter including an encoder for encoding a wideband signal as recited in claim 19 and a transmission circuit for transmitting the encoded wideband signal; and

ii) a receiver including a receiving circuit for receiving a transmitted encoded wideband signal and a decoder for decoding the received encoded wideband signal.

20

29. A cellular communication system as defined in claim 28, wherein one of said at least two paths comprises no filter for filtering the pitch codevector before supplying said pitch codevector to the pitch prediction error calculating device.

25

30. A cellular communication system as defined in claim 28, wherein said signal paths comprises a plurality of signal paths each provided with a filter

for filtering the pitch codevector before supplying said pitch codevector to the pitch prediction error calculating device of the same path.

31. A cellular communication system as defined in claim 30, wherein the filters of said plurality of paths are selected from the group consisting of low-pass and band-pass filters, and wherein said filters have different frequency responses.

32. A cellular communication system as defined in claim 28, wherein each pitch prediction error calculating device comprises:

a) a convolution unit for convolving the pitch codevector with the weighted synthesis filter impulse response signal and therefore calculating a convolved pitch codevector;

b) a pitch gain calculator for calculating a pitch gain in response to the convolved pitch codevector and the pitch search target vector;

c) an amplifier for multiplying the convolved pitch codevector by the pitch gain to thereby produce an amplified convolved pitch codevector; and

d) a combiner circuit for combining the amplified convolved pitch codevector with the pitch search target vector to thereby produce the pitch prediction error.

20

33. A cellular communication system as defined in claim 32, wherein said pitch gain calculator comprises a means for calculating said pitch gain  $b^{(j)}$  using the relation:

$b^{(j)} = x^T y^{(j)} / \|y^{(j)}\|^2$  where  $j = 0, 1, 2, \dots, K$ , and  $K$  corresponds to a number of signal paths,

and where  $x$  is said pitch search target vector, and  $y^{(l)}$  is said convolved pitch codevector.

34. A cellular communication system as defined in claim 28, wherein said pitch prediction error calculating device of each signal path comprises  
5 means for calculating an energy of the corresponding pitch prediction error, and wherein said selector comprises means for comparing the energies of said pitch prediction errors of the different signal paths and for choosing as the signal path having the lowest calculated pitch prediction error the signal path having the lowest calculated energy of the pitch prediction error.

10

35. A cellular communication system as defined in claim 32, wherein:

a) each of said filters of the plurality of signal paths is identified by a filter index;

b) said pitch codevector is identified by a pitch codebook index; and

15 c) said pitch codebook parameters comprise the filter index, the pitch codebook index and the pitch gain.

36. A cellular communication system as defined in claim 28, wherein said filter is integrated in an interpolation filter of said pitch codebook search  
20 device, said interpolation filter being used to produce a sub-sample version of said pitch codevector.

37. A cellular mobile transmitter/receiver unit comprising:

25 a) a transmitter including an encoder for encoding a wideband signal as recited in claim 19 and a transmission circuit for transmitting the encoded wideband signal; and



b) a receiver including a receiving circuit for receiving a transmitted encoded wideband signal and a decoder for decoding the received encoded wideband signal.

5 38. A cellular mobile transmitter/receiver unit as defined in claim 37, wherein one of said at least two paths comprises no filter for filtering the pitch codevector before supplying said pitch codevector to the pitch prediction error calculating device.

10 39. A cellular mobile transmitter/receiver unit as defined in claim 37, wherein said signal paths comprises a plurality of signal paths each provided with a filter for filtering the pitch codevector before supplying said pitch codevector to the pitch prediction error calculating device of the same path.

15 40. A cellular mobile transmitter/receiver unit as defined in claim 39, wherein the filters of said plurality of paths are selected from the group consisting of low-pass and band-pass filters, and wherein said filters have different frequency responses.

20 41. A cellular mobile transmitter/receiver unit as defined in claim 37, wherein each pitch prediction error calculating device comprises:

a) a convolution unit for convolving the pitch codevector with the weighted synthesis filter impulse response signal and therefore calculating a convolved pitch codevector;

25 b) a pitch gain calculator for calculating a pitch gain in response to the convolved pitch codevector and the pitch search target vector;

c) an amplifier for multiplying the convolved pitch codevector by the pitch gain to thereby produce an amplified convolved pitch codevector; and

d) a combiner circuit for combining the amplified convolved pitch codevector with the pitch search target vector to thereby produce the pitch prediction error.

42. A cellular mobile transmitter/receiver unit as defined in claim 41, wherein said pitch gain calculator comprises a means for calculating said pitch gain  $b^{(j)}$  using the relation:

$$b^{(j)} = x^T y^{(j)} / \|y^{(j)}\|^2 \quad \text{where } j = 0, 1, 2, \dots, K, \text{ and } K \text{ corresponds to a number of signal paths,}$$

and where  $x$  is said pitch search target vector, and  $y^{(j)}$  is said convolved pitch codevector.

43. A cellular mobile transmitter/receiver unit as defined in claim 37, wherein said pitch prediction error calculating device of each signal path comprises means for calculating an energy of the corresponding pitch prediction error, and wherein said selector comprises means for comparing the energies of said pitch prediction errors of the different signal paths and for choosing as the signal path having the lowest calculated pitch prediction error the signal path having the lowest calculated energy of the pitch prediction error.

44. A cellular mobile transmitter/receiver unit as defined in claim 41, wherein:

a) each of said filters of the plurality of signal paths is identified by a filter index;

b) said pitch codevector is identified by a pitch codebook index;  
and

5 c) said pitch codebook parameters comprise the filter index, the pitch codebook index and the pitch gain.

45. A cellular mobile transmitter/receiver unit as defined in claim 37, wherein said filter is integrated in an interpolation filter of said pitch codebook search device, said interpolation filter being used to produce a sub-sample version  
10 of said pitch codevector.

46. A cellular network element comprising:

a) a transmitter including an encoder for encoding a wideband signal as recited in claim 19 and a transmission circuit for transmitting the  
15 encoded wideband signal; and

b) a receiver including a receiving circuit for receiving a transmitted encoded wideband signal and a decoder for decoding the received encoded wideband signal.

20 47. A cellular network element as defined in claim 46, wherein one of said at least two paths comprises no filter for filtering the pitch codevector before supplying said pitch codevector to the pitch prediction error calculating device.

25 48. A cellular network element as defined in claim 46, wherein said signal paths comprises a plurality of signal paths each provided with a filter for

filtering the pitch codevector before supplying said pitch codevector to the pitch prediction error calculating device of the same path.

49. A cellular network element as defined in claim 48, wherein the filters of said plurality of paths are selected from the group consisting of low-pass and band-pass filters, and wherein said filters have different frequency responses.

50. A cellular network element as defined in claim 46, wherein each pitch prediction error calculating device comprises:

10 a) a convolution unit for convolving the pitch codevector with the weighted synthesis filter impulse response signal and therefore calculating a convolved pitch codevector;

b) a pitch gain calculator for calculating a pitch gain in response to the convolved pitch codevector and the pitch search target vector;

15 c) an amplifier for multiplying the convolved pitch codevector by the pitch gain to thereby produce an amplified convolved pitch codevector; and

20 d) a combiner circuit for combining the amplified convolved pitch codevector with the pitch search target vector to thereby produce the pitch prediction error.

51. A cellular network element as defined in claim 50, wherein said pitch gain calculator comprises a means for calculating said pitch gain  $b^{(j)}$  using the relation:

25 
$$b^{(j)} = x^t y^{(j)} / \|y^{(j)}\|^2 \quad \text{where } j = 0, 1, 2, \dots, K, \text{ and } K \text{ corresponds to a number of signal paths,}$$

and where  $x$  is said pitch search target vector, and  $y^{(j)}$  is said convolved pitch codevector.

52. A cellular network element as defined in claim 46, wherein said pitch  
5 prediction error calculating device of each signal path comprises means for calculating an energy of the corresponding pitch prediction error, and wherein said selector comprises means for comparing the energies of said pitch prediction errors of the different signal paths and for choosing as the  
10 signal path having the lowest calculated pitch prediction error the signal path having the lowest calculated energy of the pitch prediction error.

53. A cellular network element as defined in claim 50, wherein:

- a) each of said filters of the plurality of signal paths is identified by a filter index;
- 15 b) said pitch codevector is identified by a pitch codebook index;
- and
- c) said pitch codebook parameters comprise the filter index, the pitch codebook index and the pitch gain.

20 54. A cellular network element as defined in claim 46, wherein said filter is integrated in an interpolation filter of said pitch codebook search device, said interpolation filter being used to produce a sub-sample version of said pitch codevector.

25 55. In a cellular communication system for servicing a large geographical area divided into a plurality of cells, comprising: mobile transmitter/receiver

units; cellular base stations, respectively situated in said cells; and control terminal for controlling communication between the cellular base stations:

a bidirectional wireless communication sub-system between each mobile unit situated in one cell and the cellular base station of said one cell, said bidirectional wireless communication sub-system comprising, in both the

5 mobile unit and the cellular base station:

a) a transmitter including an encoder for encoding a wideband signal as recited in claim 19 and a transmission circuit for transmitting the encoded wideband signal; and

b) a receiver including a receiving circuit for receiving a transmitted

10 encoded wideband signal and a decoder for decoding the received encoded wideband signal.

56. A bidirectional wireless communication sub-system as defined in claim 55, wherein one of said at least two paths comprises no filter for filtering the

15 pitch codevector before supplying said pitch codevector to the pitch prediction error calculating device.

57. A bidirectional wireless communication sub-system as defined in claim 55, wherein said signal paths comprises a plurality of signal paths each

20 provided with a filter for filtering the pitch codevector before supplying said pitch codevector to the pitch prediction error calculating device of the same path.

58. A bidirectional wireless communication sub-system as defined in claim

25 57, wherein the filters of said plurality of paths are selected from the group consisting of low-pass and band-pass filters, and wherein said filters have different frequency responses.

59. A bidirectional wireless communication sub-system as defined in claim 55, wherein each pitch prediction error calculating device comprises:

a) a convolution unit for convolving the pitch codevector with the weighted synthesis filter impulse response signal and therefore calculating a convolved pitch codevector;

5                   b) a pitch gain calculator for calculating a pitch gain in response to the convolved pitch codevector and the pitch search target vector;

c) an amplifier for multiplying the convolved pitch codevector by the pitch gain to thereby produce an amplified convolved pitch codevector; and

10                   d) a combiner circuit for combining the amplified convolved pitch codevector with the pitch search target vector to thereby produce the pitch prediction error.

60. A bidirectional wireless communication sub-system as defined in claim 15 59, wherein said pitch gain calculator comprises a means for calculating said pitch gain  $b^{(j)}$  using the relation:

$$b^{(j)} = x^T y^{(j)} / \|y^{(j)}\|^2 \quad \text{where } j = 0, 1, 2, \dots, K, \text{ and } K \text{ corresponds to a number of signal paths,}$$

20

and where  $x$  is said pitch search target vector, and  $y^{(j)}$  is said convolved pitch codevector.

61. A bidirectional wireless communication sub-system as defined in claim 25 55, wherein said pitch prediction error calculating device of each signal path comprises means for calculating an energy of the corresponding pitch prediction error, and wherein said selector comprises means for comparing

the energies of said pitch prediction errors of the different signal paths and for choosing as the signal path having the lowest calculated pitch prediction error the signal path having the lowest calculated energy of the pitch prediction error.

5      62. A bidirectional wireless communication sub-system as defined in claim 59, wherein:

        a) each of said filters of the plurality of signal paths is identified by a filter index;

        b) said pitch codevector is identified by a pitch codebook index;

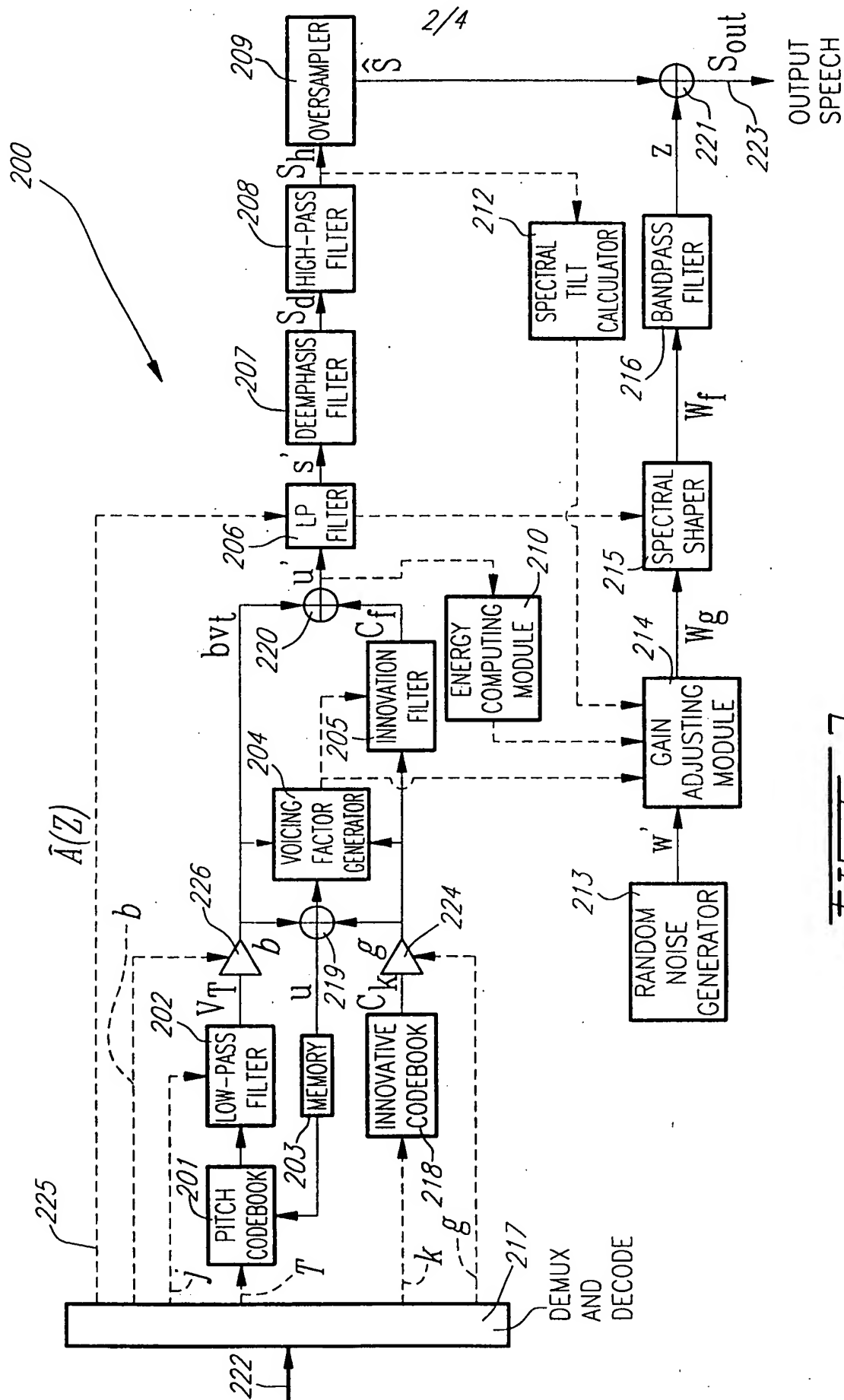
10     and

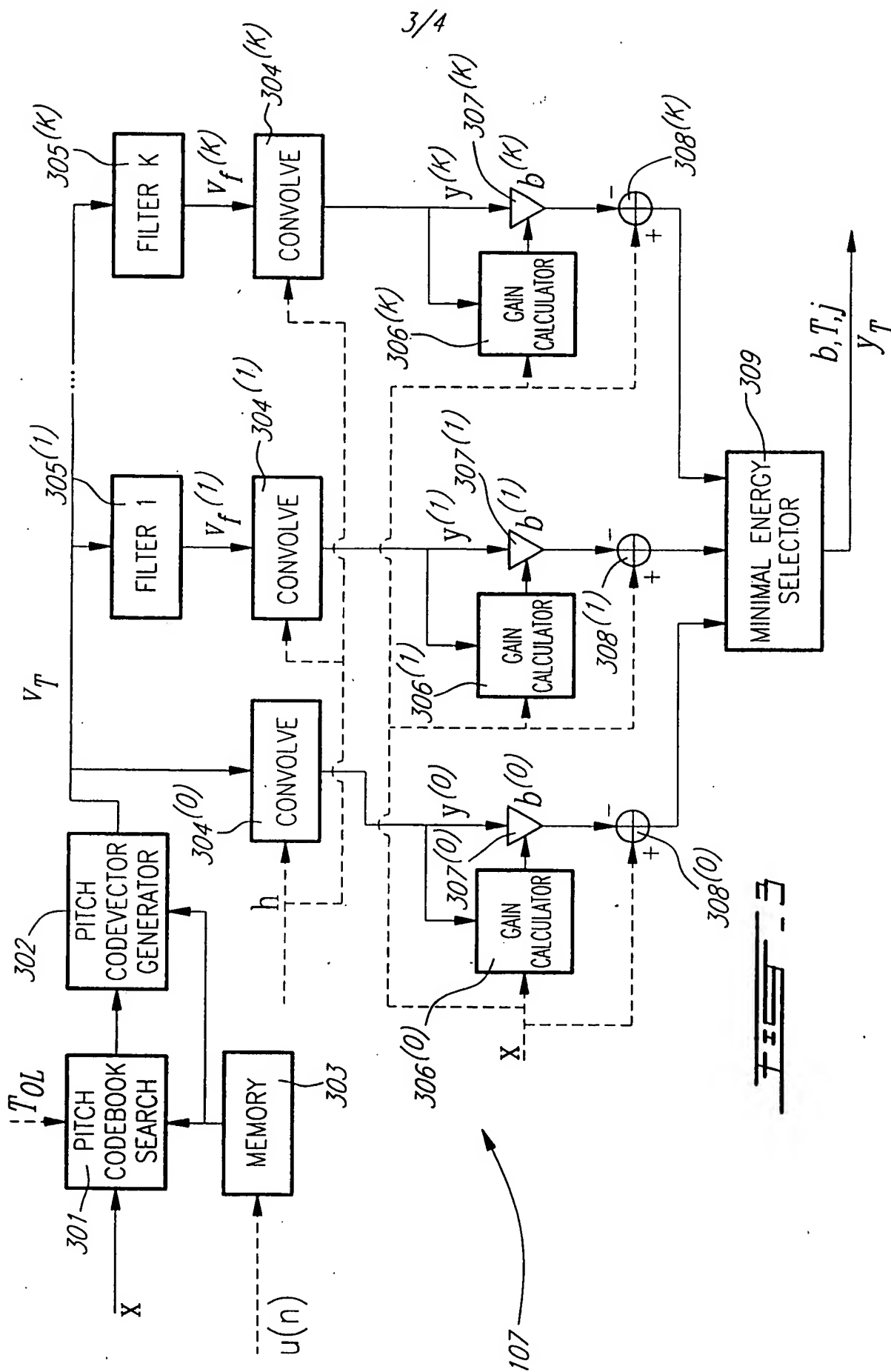
        c) said pitch codebook parameters comprise the filter index, the pitch codebook index and the pitch gain.

15     63. A bidirectional wireless communication sub-system as defined in claim 55, wherein said filter is integrated in an interpolation filter of said pitch codebook search device, said interpolation filter being used to produce a sub-sample version of said pitch codevector.

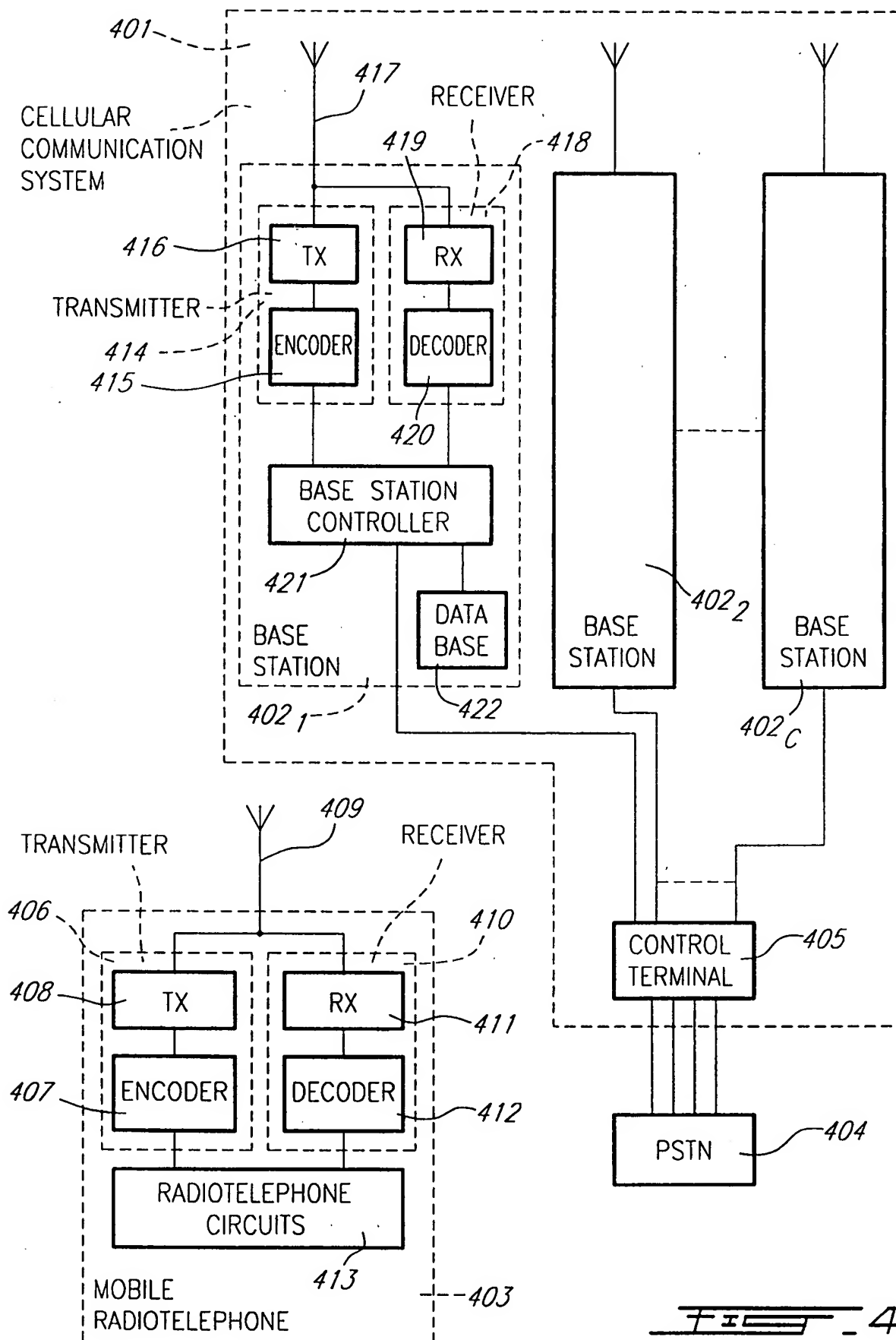








4/4

FIG. 4